The Poisson Distribution

Sam Ewen

1 Introduction

The Poisson distribution models the number of events (sometimes called arrivals) occurring in a unit time interval, where the parameter λ is the mean number of events per unit time (rate), and the time intervals between successive events are assumed to be independent of each other.

The distribution was first introduced by Simon Denis Poisson $(1781 - 1840)$. A practical application of this distribution was made by Ladislaus Bortkiewicz in 1898 when he was given the task of investigating the number of soldiers in the Prussian army killed accidentally by horse kicks; this experiment introduced the Poisson distribution to the field of reliability engineering. [1]

When our distribution matches the criteria for a Poisson Distribution, we can write $X \sim \text{Poisson}\,(\lambda)$ when X is some discrete random variable.

2 Attributes

2.1 Probability Mass Function

The probability mass function of x is given by $P\left(X = x\right) = \frac{e^{-\lambda}\lambda^{x}}{x!}$. This has been plotted in Figure 1 below. As you can see, it shows how the probabilities of given number of events are spread about the means. Three values of λ; 1, 4 and 10 are used as example cases.

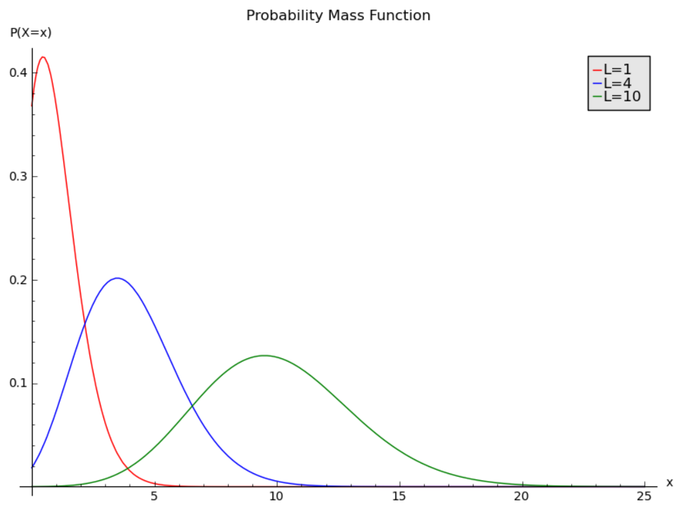

Figure 1: Plot of the PMF of the Poisson Distribution

2.2 Mean

Let $X \sim \text{Poisson}(\lambda)$

$$\begin{aligned}
\mathbb{E}(X) = \sum_{k=0}^{\infty} k\mathbb{P}(X=k) &= e^{-\lambda}\sum_{k=1}^{\infty}\frac{k\lambda^k}{k!} && \text{because the term for k = 0 is zero} \\
&= e^{-}\lambda\sum_{k=1}^{\infty}\frac{k\lambda^k}{k!}0 && \text{because } k! = k(k-1)! \\
&= \lambda e^{-}\lambda\sum_{j=0}^{\infty}\frac{\lambda^k}{j!} && \text{where we have set } j = k-1 \\
&= \lambda && \text{because} \sum_{j=0}^{\infty}\frac{\lambda^j}{j!} = e^{\lambda}
\end{aligned}$$

2.3 Variance

To find the variance, we must know $\mathbb{E}(X^2)$ and $\mathbb{E}(X)^2$. We know $\mathbb{E}(X)^2$ is just λ^2.

$$\begin{aligned}
\mathbb{E}(X^2) = \sum_{k=0}^{\infty} k^2\mathbb{P}(X=k) &= e^{-}\lambda\sum_{k=0}^{\infty}\frac{k^2\lambda^k}{k!} \\
&= e^{-}\lambda\sum_{k=1}^{\infty}\frac{k^2\lambda^k}{k!} && \text{because the term for } k = 0 \text{ is zero} \\
&= \lambda e^{-}\lambda\sum_{k=1}^{\infty}\frac{k\lambda^k - 1}{(k-1)!} && \text{because } k! = k(k-1)! \\
&= \lambda e^{-}\lambda\sum_{j=0}^{\infty} && \text{where we have set } j = k-1 \\
&= \lambda e^{-}\lambda[\sum_{j=0}^{\infty}\frac{j\lambda^j}{j!} + \sum_{j=0}^{\infty}\frac{\lambda^j}{j!}] \\
&= \lambda e^{-}\lambda[\lambda e^{\lambda} + e^{\lambda}] && \text{because } e^{-}\lambda\sum_{j=0}^{\infty}\frac{\lambda^j}{j!} = \mathbb{E}(X) = \lambda \\
&= \lambda(\lambda + 1)
\end{aligned}$$

$$\therefore \text{Var}(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2 = \lambda(\lambda+1) - \lambda^2 = \lambda$$

So after finding these two results we can conclude that in the Poisson Distribution, the mean and the variance obtain the same value.

3 Poisson Approximation of the Binomial Distribution

We can use Poisson to give an approximation of the Binomial distribution in particular circumstances. If we take the parameter $\lambda = np$, and n is large and p is small, we can use the Poisson distribution as an approximation to $B(n, p)$ of the binomial distribution. Obviously it is a matter of opinion and relation whether a value is 'large' or 'small'; general guidelines are that this approximation is good if $n \geq 20$ and $p \leq 0.05$, or if $n \geq 100$ and $np \leq 10$ [2]

For example, if you take the number of trials(n) to be 160, and the probability of success(p) to be 0.05, then we can use the Poisson distribution to approximate this spread as shown in Figure 2. We do this by letting $np = \lambda$, so in our case that makes our λ equal to 8.

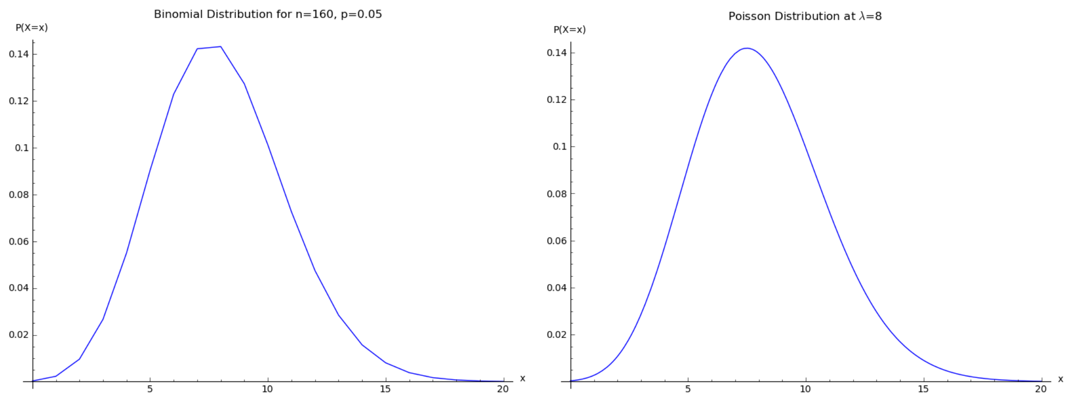

Figure 2: Approximation for Binomial Distribution $n = 160$, $p = 0.05$ using Poisson Distribution $\lambda = 8$

SAGE can be used to create these plots, but different techniques must be used to accomodate for the varied input of the two distributions. For the Binomial Distribution, a 'for' loop can be used to go through the formula for all desired values, and append the results to a list. For this it is necessary to input values for n and p. Using this method, in the form of the code below, we can create the plot.

```
def Bin(n,p):
    val = []
    for k in range(n):
        nCk = factorial(n)/(factorial(k)*factorial(n-k))
        d = nCk*(p^k)*(1-p)^(n-k)
        val.append(d)
    return list_plot(val, plotjoined=True, axes_labels=['x', 'P(X=x)']
```

To plot the Poisson Distribution for a given value of λ, the following code can be used:

```
L = var('L')
f(x,L) = ((L^x)*(e^(-L)))/factorial(x)
'''
Define function:

Function that takes L as an abbreviation for lambda, and returns the Probability Mass Function(PMF).
Input: 3 different values of lambda (means) into the equation
for the PMF of the Poisson Distribution.
Output: A spread of probabilities showing how likely each value is to occur around the given mean L.

'''
```

References

[1] Wikipedia. Poisson distribution — wikipedia, the free encyclopedia, 2013. [Online; accessed 9-December-2013].

[2] Wikipedia. Binomial distribution — wikipedia, the free encyclopedia, 2013. [Online; accessed 10-December-2013].